An Analysis of the Effect of Global Warming on the Intensity of Atlantic Hurricanes Using a GCM with Statistical Refinement

MING ZHAO

University Corporation for Atmospheric Research, Boulder, Colorado, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

ISAAC M. HELD

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 10 May 2010, in final form 9 August 2010)

ABSTRACT

A statistical intensity adjustment is utilized to extract information from tropical cyclone simulations in a 50-km-resolution global model. A simple adjustment based on the modeled and observed probability distribution of storm lifetime maximum wind speed allows the model to capture the differences between observed intensity distributions in active/inactive year composites from the 1981–2008 period in the North Atlantic. This intensity adjustment is then used to examine the atmospheric model's responses to different sea surface temperature anomalies generated by coupled models for the late twenty-first century. In the North Atlantic all simulations produce a reduction in the total number of cyclones, but with large intermodel spread in the magnitude of the reduction. The intensity response is positively correlated with changes in frequency across the ensemble. However, there is, on average, an increase in intensity in these simulations despite the mean reduction in frequency. The authors argue that it is useful to decompose these intensity changes into two parts: an increase in intensity that is intrinsic to the climate change experiments and a change in intensity positively correlated with frequency, just as in the active/inactive historical composites. By isolating the intrinsic component, which is relatively independent of the details of the SST warming pattern, an increase is found in storm-lifetime maximum winds of 5–10 m s^{-1} for storms with intensities of 30–60 m s^{-1}, by the end of the twenty-first century. The effects of change in frequency, which are dependent on the details of the spatial structure of the warming, must then be superimposed on this intrinsic change.

1. Introduction

Several recent studies suggest that general circulation models (GCMs) with resolutions in the range of 20–100 km are adequate for studying many aspects of tropical cyclone (TC) genesis and TC frequency distributions (e.g., Bengtsson et al. 2007a,b; Oouchi et al. 2006; LaRow et al. 2008; Zhao et al. 2009, hereinafter ZHLV; Vitart 2006; Gualdi et al. 2008). However, these models are typically unable to simulate a realistic frequency of TCs in the higher intensity categories that account for most storm damage. One approach to correcting this deficiency is to use an additional dynamical downscaling step, as in Bender et al. (2010). In the present study we consider an alternative simple statistical refinement approach. Using this statistical adjustment, we investigate the ability of a GCM with roughly 50-km resolution to simulate aspects of the interannual variability of TC intensities in the North Atlantic. Encouraged by these results, we also estimate the response of TC intensity distribution to global warming. In the process, we contrast the relationship between TC frequency and intensity in interannual variations and in the simulated response to global warming.

The model we use here is the Geophysical Fluid Dynamics Laboratory (GFDL) High Resolution Atmospheric Model (HIRAM), specifically C180HIRAM2.1 [C180 indicates a cubed sphere dynamical core (Putman and Lin 2007) with 180 × 180 grid points on each face of the cube, resulting in grid sizes ranging from 43.5 to 61.6 km]. The climatological statistics of hurricanes simulated in this model, as well as the interannual variability in the model when run over observed sea surface temperatures (SSTs) from the Hadley Centre Global Sea Ice

Corresponding author address: Dr. Ming Zhao, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton University Forrestal Campus, 201 Forrestal Road, Princeton, NJ 08540–6649.
E-mail: ming.zhao@noaa.gov

DOI: 10.1175/2010JCLI3837.1

© 2010 American Meteorological Society

and Sea Surface Temperature (HadISST) dataset, version 1.1[1] (Rayner et al. 2003), have been described in ZHLV. ZHLV demonstrate that the simulated hurricane statistics from a four-member ensemble compare well with the observations for the period of 1981–2005. These include the geographical distribution of storm genesis locations, as well as the seasonal cycle and the interannual variability of hurricane frequency for the North Atlantic and the east and west Pacific. As described in Zhao et al. (2010), we have continued our four-member integrations up to and including 2008 and the simulations continue to maintain their high quality in simulating hurricane frequency. The observations used in ZHLV and in Zhao et al. (2010) are taken from the International Best Track Archive for Climate Stewardship (IBTrACS; Kruk et al. 2010; http://www.ncdc.noaa.gov/oa/ibtracs/).

Despite its quality in simulating hurricane climatology and variability, the model cannot produce major Atlantic hurricanes (category 3 and higher, with maximum wind speed greater than 50 m s^{-1}), as can be seen in Fig. 6a in ZHLV and Fig. 1 in this study. While this unrealistic aspect of the intensity distribution prevents us from taking the model response of TC intensity to global warming at face value, some of the most basic differences in intensity distributions between basins are captured qualitatively, such as the presence of more intense TCs in the west Pacific as compared to the Atlantic (Figs. 6a,c in ZHLV). In this study, we use the same set of simulations (with the extension up to 2008) to explore whether we can extract useful information from the model with regard to the TC intensity distribution, including that of major hurricanes. To achieve this, we introduce a simple statistical refinement of the GCM results based only on the observed and modeled climatological intensity distribution.

Seasonal measures of TC intensity are positively correlated with TC frequency in the North Atlantic. Consistently, one also sees an increase in average intensity during La Niña as compared to El Niño years, and one finds a trend toward higher intensities over the past three decades accompanying a trend toward increased number of TCs. We use these interrelated measures to test the results obtained from the statistically adjusted GCM results. Section 2 presents the statistical intensity adjustment and then describes these comparisons with observations. Section 3 provides the results for global warming experiments using SSTs from a number of coupled model projections of the late twenty-first century. SST anomalies based on the A1B scenario are used for this purpose. Section 4 gives a summary and discussion of the intensity response to global warming, linking to the potential intensity (PI) theory of Emanuel (1988).

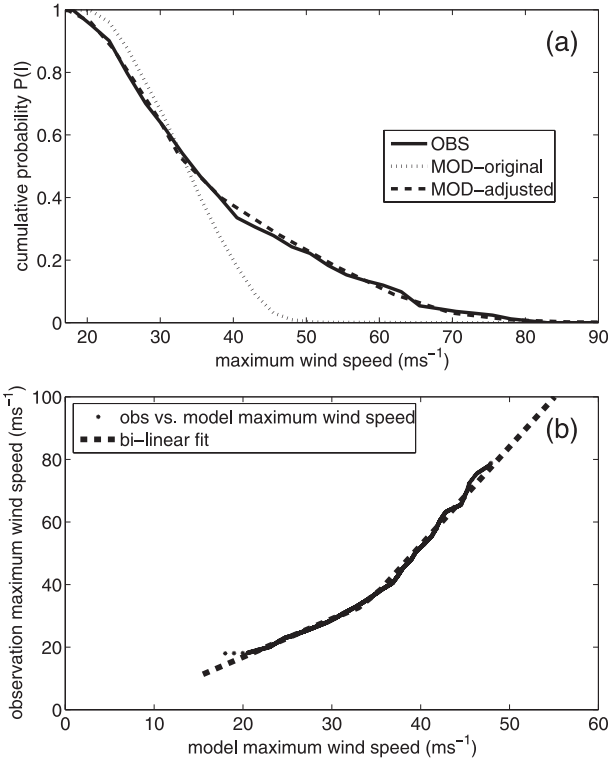

FIG. 1. (a) Cumulative probability $P(I)$ of TC lifetime maximum wind speed in the North Atlantic for 1981–2008 for both observations (solid) and the model (dotted). Dashed curve shows the model results after a statistical refinement of model wind speed based on the bilinear relationship shown in (b). (b) Plot of TC lifetime maximum wind speed between the model and the observations based on equal value of $P(I)$ shown in (a). Straight lines are two linear fits, one for tropical storm regime and the other for hurricane regime.

2. The statistical refinement and results for 1981–2008

We define I to be the maximum near-surface wind speed over the lifetime of a tropical cyclone (TCs are here defined as any cyclones with $I \geq 17$ m s^{-1} and the model's TC identification algorithm is discussed in ZHLV) and $F(I)$ as the total number of TCs per year with intensity exceeding I. We also write $F(I) = NP(I)$, where N denotes the total number of TCs and $P(I)$ denotes the probability that a given TC exceeds the intensity I. Since $\delta F(I) \approx P(I)\delta N + N\delta P(I)$ for small perturbations to the climate, to realistically simulate $\delta F(I)$ a model needs to

[1] The model-simulated climatology of North Atlantic hurricanes using the NOAA optimum interpolation SST analysis dataset (Reynolds et al. 2002) shows a (20%–30%) lower annual count of hurricanes for this simulation period (ZHLV). A systematic study on the model sensitivity to the different SST products using the long-term record will be reported elsewhere (G. Vecchi et al. 2010, unpublished manuscript).

produce not only realistic δN but also a realistic $P(I)$ and its variation $\delta P(I)$. As seen in ZHLV, a 50-km GCM can simulate observed δN reasonably well, using interannual variability and decadal trends as a measure of quality, even though it generates an unrealistic $P(I)$ for large I. Below we further illustrate this problem and then introduce a very simple statistical refinement to correct this model bias.

Figure 1a shows both the modeled and the observed $P(I)$ in the North Atlantic for the 1981–2008 period. While the model captures reasonably well the cumulative probability for TCs with intensity up to the hurricane threshold ($I \leq 33$ m s^{-1}), beyond this intensity the probability drops off rapidly as compared with observations. Category 3 and higher storms ($I \geq 50$ m s^{-1}) are missing entirely in the model. The observed $P(I)$ flattens above the hurricane threshold [as discussed by Emanuel (2000) and Swanson (2007)], while the model generates about the same slope beyond the hurricane criterion as for lower intensities.

It would clearly be beneficial if this systematic bias in $P(I)$ could be corrected in a way that would allow the extraction of significant information concerning major hurricane frequency. To do this we first identify the modeled and observed values of I corresponding to equal values of P. Fig. 1b shows a plot of the modeled and observed maximum wind speed from this equal-P match. The data points fall along a simple curve that can be well represented by two straight lines with an intersection at a model wind speed close to the hurricane threshold:

$$I_{\mathrm{fit}} = I_T + \alpha(I_{\mathrm{GCM}} - I_T); \quad I_{\mathrm{GCM}} \leq I_T, \tag{1a}$$

$$I_{\mathrm{fit}} = I_T + \beta(I_{\mathrm{GCM}} - I_T); \quad I_{\mathrm{GCM}} > I_T, \tag{1b}$$

with $\alpha = 1.22$, $\beta = 3.15$, and $I_T = 34$ m s^{-1}.

The use of two separate linear fits (one for hurricanes and the other for tropical storms) is consistent with an earlier study on TC intensity distribution (Emanuel 2000) as well as with the strong impression of a distinctive model deficiency that sets in for maximum winds greater than hurricane strength. This simple bilinear relationship forms the basis for our statistical refinement; we simply use these expressions to convert the GCM intensity into an adjusted intensity estimate. The dashed curve in Fig. 1a shows the model $P(I)$ after this statistical correction. This correction leaves the number of hurricanes essentially unchanged and reduces slightly the total number of TCs.

To see if this correction adds value to the model, we examine whether the resulting adjusted model can simulate differences in intensities in the observational record from 1981 to 2008. It is known that TCs in active years in the North Atlantic tend to have higher intensities on average than TCs in inactive years (e.g., Swanson 2007). A year-by-year comparison of observed and modeled intensities is too noisy to be directly useful. Therefore, we consider three alternative composites to quantify this relationship: the 10 most active years versus the 10 least active years in this period as measured by number of TCs (see appendix for a list of years); La Niña years versus El Niño years; and the last half of this period (1995–2008) versus the first half (1981–94), there being an upward trend in intensity and frequency over this period. The three tests are not independent, but the alternative groupings of years provide increased confidence that the model–observation comparison is robust.

Figure 2 shows that the raw GCM output provides a reasonably good estimate for the difference in the average number of TCs per year, and the number of hurricanes, in each of these three composite pairs. The change in total number of TCs is underestimated for two of the three composite pairs, while the hurricane count change is significantly underestimated in only one of the three. For major hurricanes, all of which are generated by the statistical adjustment, the magnitude of the change in number is captured fairly well, once again, with the change being underestimated in two of the three composites.

The figure also shows the change in mean intensity, I_0, obtained by averaging the maximum intensity of each storm over all TCs in given years. The raw GCM simulation of this intensity change has the correct sign but is far too small in amplitude. The statistical correction provides much improved estimates of this average intensity shift, although still somewhat smaller than observed in two of the three composites.

Figure 3 provides more detail on the change in intensity distribution in these composites. We define the change in intensity δI at a given intensity I by first computing the probability of exceedance $P(I)$ for the active and inactive cases and converting this into the intensity change at each value of P. We then convert this from a function of P into a function of I using the climatological $P(I)$ distribution. The observed changes ramp up to values of about 10 m s^{-1} for values of I near 50 m s^{-1}, and then level off for even stronger storms. The values in the unadjusted model bear little resemblance to the observations. The adjustment stretches the abscissa by a factor of 3.15 for hurricane intensities and the ordinate by roughly the same factor (the ordinate would be stretched by exactly the same factor if the active/inactive difference were infinitesimal). The adjustment produces curves of $\delta I(I)$ with shapes comparable to that observed. Typically, the resulting changes in intensity at values greater that 50 m s^{-1} are roughly 15%–30% smaller than the estimates from observations.

For a slightly different perspective, we also plot in the same figure the change in the probability of exceedance

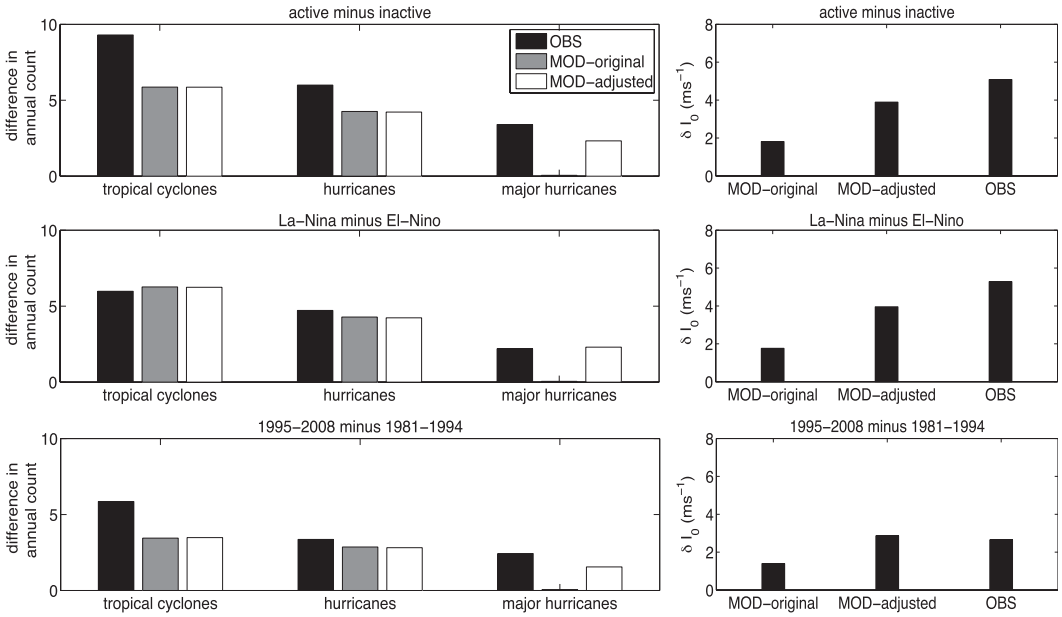

FIG. 2. (left) Model and observed difference in annual count of all TCs, hurricanes, and major hurricanes between (top) the active and inactive years, (middle) the La Niña years and El Niño years, and (bottom) the 1995–2008 and 1981–94 periods. (right) As at left, but for the difference in TC mean intensity.

per unit change in TC count, $\delta P/\delta N$. After statistical adjustment the fits are equally reasonable at all hurricane intensities in each of the three cases. The observations show a structure that is relatively flat between 30 and 60 m s^{-1}, with an increase in probability of about 0.02 for each increase of one TC per season. The adjusted model produces similar shapes and similar amplitudes. The impression that the model underpredicts the response is less evident here than in the plots of $\delta I(I)$ because of the normalization by δN. This normalized probability change will prove to be of interest in the discussion of our global warming simulations.

Since $F(I) \equiv NP(I)$, using centered differencing with respect to a change of large-scale climate condition (e.g., SSTs) we can decompose $\delta F/\delta N$ between active and inactive years into two parts, corresponding to the change in overall number and the change in the intensity distribution:

$$\frac{\delta F}{\delta N} = \tilde{P} + \tilde{N}\frac{\delta P}{\delta N}, \tag{2}$$

where $\delta N \equiv N_+ - N_-$, $\tilde{N} = (N_+ + N_-)/2$, and similarly for P. (The centered differencing avoids the need for considering any additional quadratic terms in this expansion.) Figure 4a shows this decomposition for one of our composites: the difference between the 10 most active and least active years. The normalization by δN does not change the relative magnitude of these two terms, but it does force the model and the observations to agree at the low intensity limit. This allows the model error in simulating the total TC frequency variability (which is independent from our statistical adjustment) to be factored out in this comparison. The relative magnitude of these contributions is very similar in the observations and the adjusted model results, with the change in overall number being dominant for $I < 40$ m s^{-1} but with the change in intensity distribution growing to comparable or even larger values for $I > 55$ m s^{-1}.

Qualitatively similar results are obtained for the east Pacific as for the North Atlantic, although the required bilinear fit differs quantitatively from that optimized for the North Atlantic. We do not discuss these east Pacific results here. The results for the west Pacific are not as impressive, despite the fact that there is documentation of an El Niño signal in TC intensities in that basin (Camargo and Sobel 2005).

Evidently, the statistical adjustment is more than a cosmetic way of improving the model's intensity distribution; it also allows us to extract useful information as to the sensitivity of the intensity distribution to the changes in SSTs associated with interannual variability. It generates model simulations for the differences in intensity distribution between active and inactive years that compare favorably to the observed changes, although these differences tend to be somewhat smaller than observed. This comparison with observations suggests that it is useful to examine the response of the same model to global

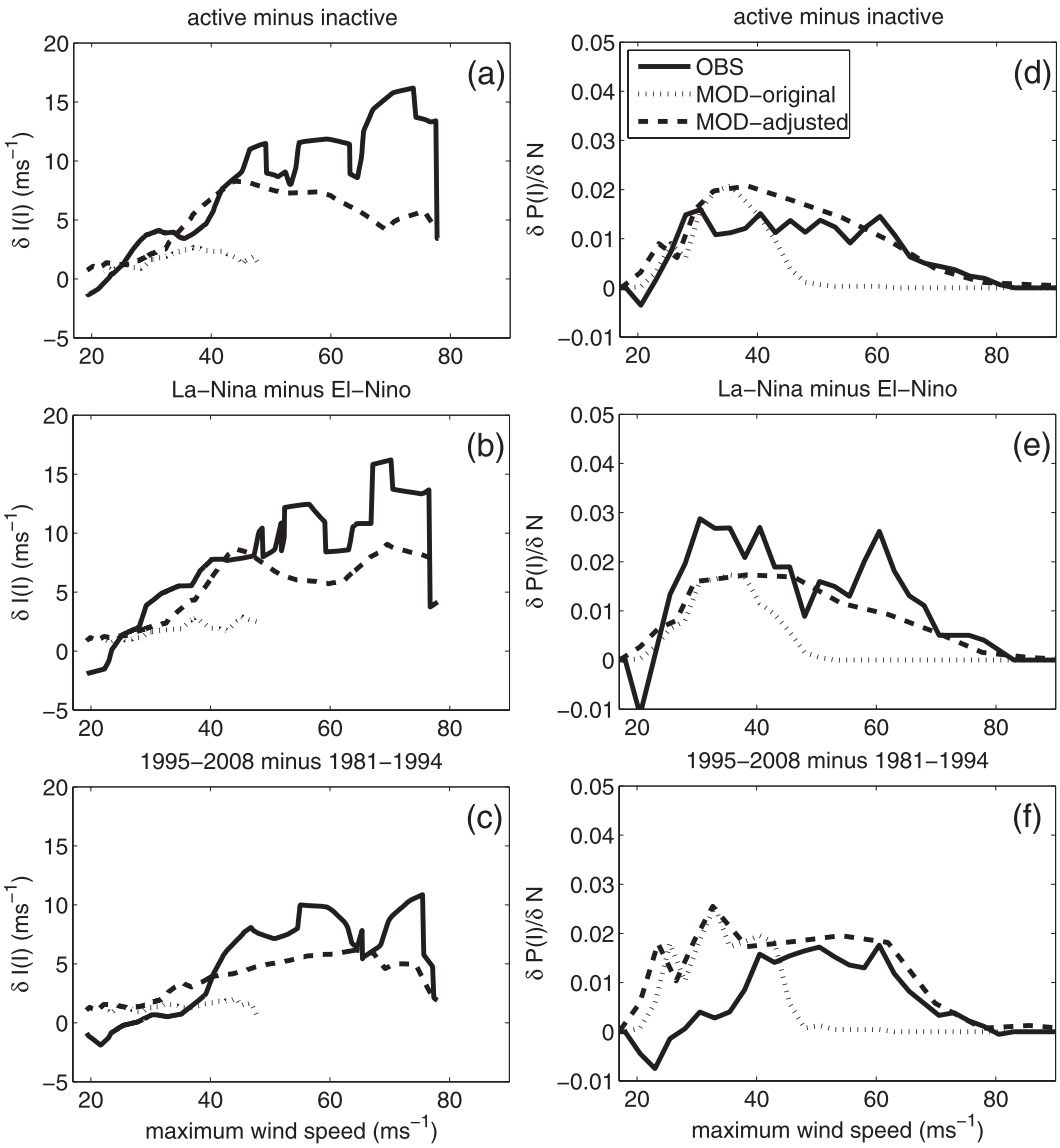

FIG. 3. (left) As in Fig. 2, but for model and observed differences in TC intensity $\delta I(I)$. (right) As at left, but for $\delta P(I)/\delta N$.

warming, using the identical statistical intensity adjustment.

3. Global warming experiments

In ZHLV, the identical model was used to simulate the hurricane response to four different SST anomalies generated by coupled models in the archive of phase 3 of the Coupled Model Intercomparison Project (CMIP3; https://esg.llnl.gov:8443/index.jsp) (Meehl et al. 2007) for the late twenty-first century based on the A1B scenario. The SST anomalies were obtained from single realizations of three models [GFDL's Climate Model version 2.1 (CM2.1), the third climate configuration of the Met Office Unified Model (HADCM3), and the Max Planck Institute's ECHAM5)], and from the ensemble mean for the simulations for 18 models. Since the ZHLV study, we have further pursued five additional SST warming experiments. Four of these are analogous to those in ZHLV but use SST anomalies from different coupled models [GFDL CM2.0, the Hadley Centre Global Environmental Model version 1 (HADGEM1), the Meteorological Research Institute (MRI) coupled general circulation model version 2.3.2 (CGCM2.3.2), and the Center for Climate System Research (CCSR) Model for Interdisciplinary Research on Climate, high-resolution version (MIROC-HI)]. In

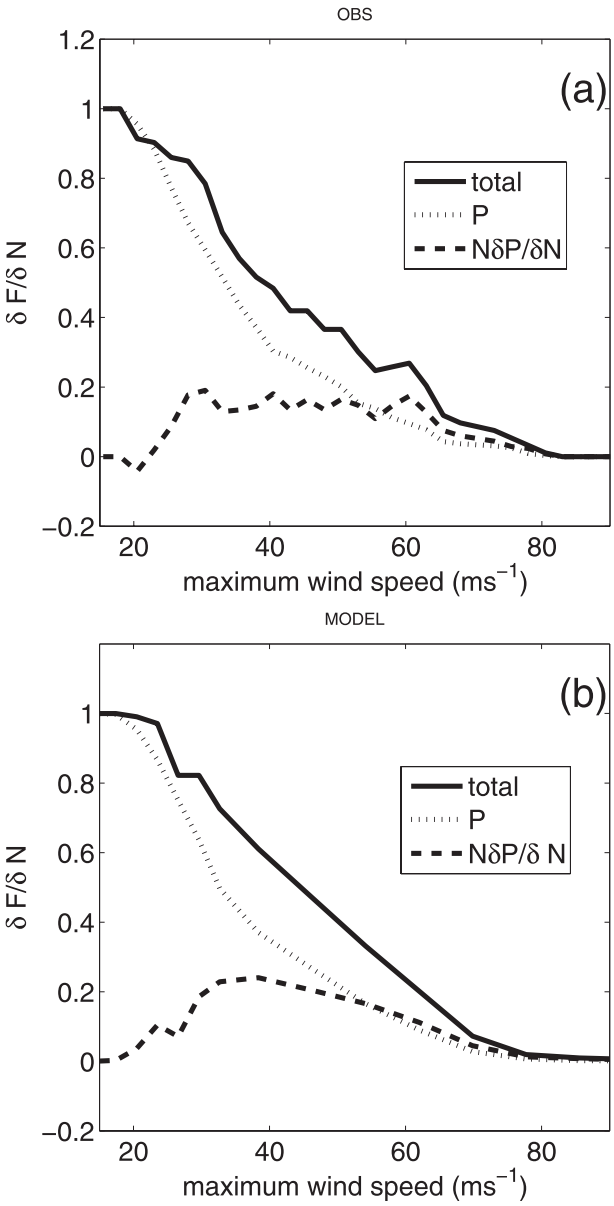

FIG. 4. Decomposition of $\delta F(I)/\delta N$ (solid) into the two terms $\tilde{P}(I)$ (dotted) and $\tilde{N}\delta P(I)/\delta N$ (dashed) for the (top) observation and (bottom) model, using the difference between the 10 most and least active years. See Eq. (2) for the decomposition and the definition of F, $\tilde{P}$, and $\tilde{N}$.

addition to these coupled model projections, we have also performed a simple experiment with the control SST uniformly warmed by 2 K (called P2K below). As described in ZHLV, we compute the multimodel ensemble mean SST anomaly by differencing the period 2081–2100 and the period 2001–20 from the A1B simulations in the CMIP3 archive. For each of the seven individual models, we use one realization (run 1 in the CMIP3 archive) to compute the 2001–2100 linear trend. The linear trend is then multiplied by 0.8 so that it is consistent in magnitude with the 80-yr period (2010–90) used for the multimodel ensemble mean. The use of century-scale linear trend instead of the difference between the two end periods helps reduce the contribution of internal variability in the individual models since only one realization is used for each model.

As described in ZHLV, we have generated two control simulations by prescribing the climatological SSTs (seasonally varying with no interannual variability) using the time-averaged HadISST dataset (Rayner et al. 2003) as well as the National Oceanic and Atmospheric Administration (NOAA) optimum interpolation SST analysis dataset (Reynolds et al. 2002). We then add the SST warming anomalies (also seasonally varying with no interannual variability) projected by the coupled models to each of the climatological SSTs to pursue the global warming experiments. For the five additional global warming experiments since the ZHLV study, we have only conducted the warming experiments with respect to the HadISST control climatological SST. In all cases, including P2K, we also double the concentration of CO_2 in the atmosphere, a change roughly comparable to that imposed over the twenty-first century in the A1B experiments. As in ZHLV, 10-yr integrations were carried out for both the controls and the perturbation experiments. To be sure that 10-yr statistics are sufficient for storm intensity analysis, we have also selectively extended a few experiments (the control, 18-model ensemble mean warming, P2K, and GFDL CM2.1 for the HadISST background climatology) to 20 years. We do not find substantial differences between the 10-yr and 20-yr statistics. We use the 20-yr averages and the averages over the two background climatology experiments for the cases available in the results described below. We then apply the statistical intensity correction described in section 2 to these global warming experiments to investigate storm responses to twenty-first-century warming.

Before displaying the full range of storm intensity response and its decomposition, Fig. 5 shows a scatterplot of the changes in annual count of Atlantic TCs, hurricanes, and major hurricanes in all of the integrations versus the changes in an Atlantic relative SST index. This index (below called Φ_{NA}) is defined as the August–October seasonal mean difference between the SSTs averaged over the North Atlantic Main Development Region (MDR) (10°–25°N, 80°–20°W) and the SSTs averaged over the entire tropics (30°S–30°N) (e.g., Swanson 2008; Vecchi et al. 2008; ZHLV). ZHLV (their Fig. 16) show that the Φ_{NA} anomaly explains much of the spread in the four global warming simulations using the SST anomalies projected by global coupled models. Zhao et al. (2010) (their Fig. 5a) further demonstrate that the index is well

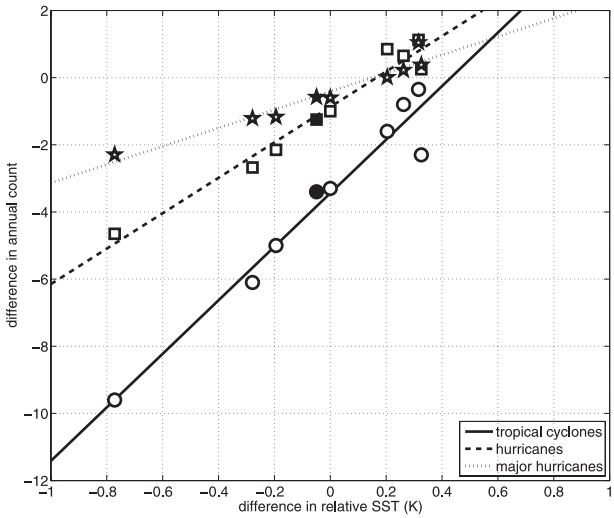

FIG. 5. Scatterplots of the difference (warming minus control) in North Atlantic Main Development Region (MDR) relative SST and the corresponding difference in the annual count of Atlantic TCs (circles), hurricanes (squares), and major hurricanes (pentagrams). Filled symbols denote the results from the 18-model ensemble mean SST anomaly experiments. The results lying on the zero relative SST line are for the uniform warming (P2K) experiment. Lines are linear regressions. Relative SST is defined as the August–October seasonal mean difference between the SSTs averaged over the North Atlantic MDR and the SSTs averaged over the entire tropical ocean (30°S–30°N).

correlated interannually with both the observed ($r = 0.76$) and model-simulated ($r = 0.83$) hurricane counts and point out its implications for hurricane seasonal prediction. With five additional global warming patterns and the statistical refinement, Fig. 5 emphasizes that this index is separately well correlated with the number of TCs, hurricanes, and major hurricanes generated by the model running over these various SST warming patterns. The linear regressions for each classification give different slopes, with weaker storms having larger slopes simply because there are more such storms. More interestingly, the value of $\delta\Phi_{NA}$ that produces zero change in frequency is lower for hurricanes (0.2°C) than for all TCs (0.45°C), so that a number of simulations produce an increase in hurricane counts, but none produce an increase in the total of all TCs. While one might expect this trend to be continued for major hurricanes as compared to hurricanes in general, this is not the case; the crossing point for major hurricanes and all hurricanes are nearly the same in these results. The extent to which this is a limitation of the statistical intensity adjustment strategy is unclear.

An examination of the large-scale dynamical and thermodynamical field in the MDR indicates that the simulated intermodel spread in storm frequency response can also be well explained by both potential intensity and vertical wind shear (similar to the results in ZHLV). In both cases, the correlations are above 0.8. Also, in the case of uniform warming (i.e., P2K), the MDR vertical wind shear increases slightly, consistent with the storm reduction shown in Fig. 5. This result indicates that in addition to the change of relative SST, which largely determines the response of storm genesis frequency, uniform warming acts to suppress the Atlantic storm genesis by a modest amount in this model.

Among the different SST anomaly experiments, the result that stands out with a value of $\delta\Phi_{NA}$ close to -0.8°C and producing the largest reduction in total storm counts is due to the SSTs from HADGEM1. The North Atlantic TC count drops off from 10.9 to 1.4 storms in this case. Since the Atlantic storms in the HADGEM1 projected twenty-first-century warming experiment are so infrequent, even 20-yr simulation is insufficient to obtain stable statistics, especially for intense hurricanes. Therefore, we will exclude this model in the following analysis of storm intensity distribution, keeping in mind the presence of this distinctive outlier in the ensemble.

To examine storm response over the full range of intensities, Fig. 6a displays changes in cumulative frequency $\delta F(I)$ of the North Atlantic TCs between the global warming and the control experiments. [Since we care about the overall change of frequency over the full range of intensity it is desirable not to normalize $\delta F(I)$ by δN. Indeed, as we will show below, δN is an important component in determining the overall frequency change even for fairly intense hurricanes.] As shown in Fig. 5, all experiments produce a decline of the total number of storms, with a large spread in the magnitude of the decline. As intensity increases, the magnitude of reduction in TC activity decreases, with models producing both rise and fall for hurricanes and for major hurricanes. [Note that the seemingly narrowing spread at higher intensity does not indicate better agreement among models; it is simply due to the distribution of TC cumulative probability $P(I)$, which declines sharply with intensity. When measured by fractional change, the model spread actually slightly increases with intensity (not shown, being quite noisy)].

Both the ensemble mean of the SST anomaly experiments and the result from the 18-model ensemble mean SST anomaly show significant reductions of storms at intensities below 60 m s^{-1} with total loss of annual count of roughly 3–4 (~30%–40%). There is, however, a significant difference between the two, with the 18-model ensemble mean SST anomaly experiment producing a larger reduction. This is likely due to sampling difference. In particular, the SST projection from the HADGEM1 model is used in generating the 18-model ensemble mean SSTs but is excluded in the intensity analysis and therefore the ensemble mean of different SST warming experiments.

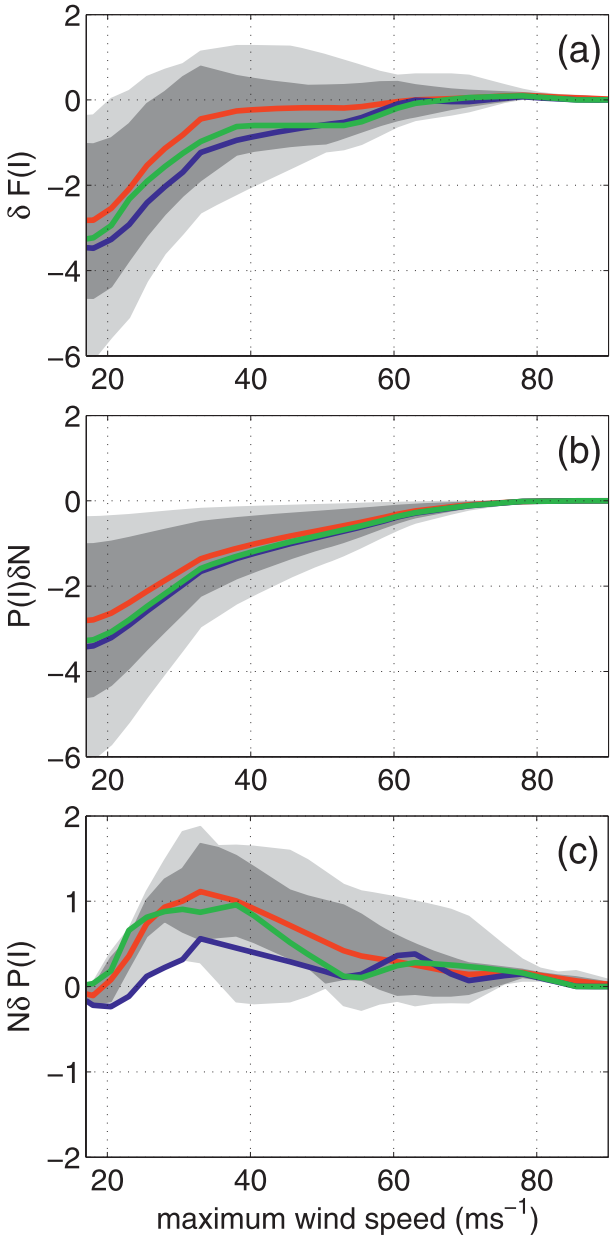

FIG. 6. (a) Difference (warming minus control) in cumulative frequency $\delta F(I)$ of North Atlantic TCs between global warming experiments and the control experiment. Red indicates the ensemble mean for the SST anomaly experiments; blue, the result for the 18-model ensemble mean SST anomaly experiment; and green, the result from the 2-K uniform warming experiment. Dark shading indicates the central 50% range (25%–75%) and light shading the minimum and maximum from the SST anomaly experiments. (b),(c) As in (a), but for $P(I)\delta N$ and $N\delta P(I)$, respectively.

More generally, the Atlantic MDR relative SST from the 18-model ensemble mean is lower (−0.15 K) than that averaged across the different SST warming experiments we have utilized, consistent with its lower Atlantic TC activity across much of the intensity spectrum. At intensities above 60 m s^{-1} (category 4 and higher), both the mean of the SST anomaly experiments and the experiment with the ensemble mean SST anomaly display only a small change, with about a half of the models generating increasing/decreasing storm activity. At intensities above 75 m s^{-1}, the results are noisy because of the small number of events, but seven out of the eight experiments produce an increase.

To better elucidate the intensity responses, Figs. 6b and 6c display the partition of $\delta F(I)$ into terms due to the changes in δN and δP: $\delta F = P\delta N + N\delta P + \delta N\delta P$. Here we use one-sided differences (perturbed minus control) so the quadratic term needs to be considered, but in fact it turns out to be negligible. Both the $P(I)\delta N$ and $N\delta P(I)$ terms contribute to the total change. Similar to the active/inactive composites in the historical experiments (see Fig. 4), $P(I)\delta N$ dominates $\delta F(I)$ at the lower range of intensity while at higher intensity the $N\delta P(I)$ term becomes a larger term. However, there is a fundamental difference here in that the global warming experiments produce increasing $P(I)$ despite declining N (Figs. 6b,c), whereas the historical active/inactive composites show decreasing $P(I)$ associated with declining N (Fig. 3, right panels).

This distinction is qualitatively consistent with recent studies (e.g., Bengtsson et al. 2007b; Bender et al. 2010) as summarized in Knutson et al. (2010), which suggest that warmer climate likely leads to more very intense Atlantic hurricanes but with an overall decrease in frequency. However, the quantitative prediction of the transition point from $\delta F(I) < 0$ to $\delta F(I) > 0$ appears to be uncertain and sensitive to both SSTs and individual downscaling models used. For example, Bender et al. (2010) used a dynamical downscaling model and found that three out of the five SST projections (including the 18-model ensemble mean SSTs) produce large reductions of intense/major (category 3 and above) hurricanes while the other two produce an increase. For category 4 and 5 hurricanes, four out of the five SST warming experiments produce an increase with one still producing a decrease. For the ensemble mean SSTs, the results of Bender et al. (2010) indicate a transition point at intensity somewhere between categories 3 and 4 (50–60 m s^{-1}). In our study the transition point appears to occur at a higher intensity (60–70 m s^{-1}) for the ensemble mean SSTs. On the other hand, our results for the ECHAM5 projected SST produce a much lower transition point than that from the ensemble mean SST, with increases of both hurricanes and intense hurricanes. This is consistent with Bengtsson et al. (2007b), who used a similar SST projection but a different global high-resolution atmospheric model. We emphasize that because of the potential for cancellation between the two tendencies (i.e., δN and δP) in a greenhouse gas–warmed

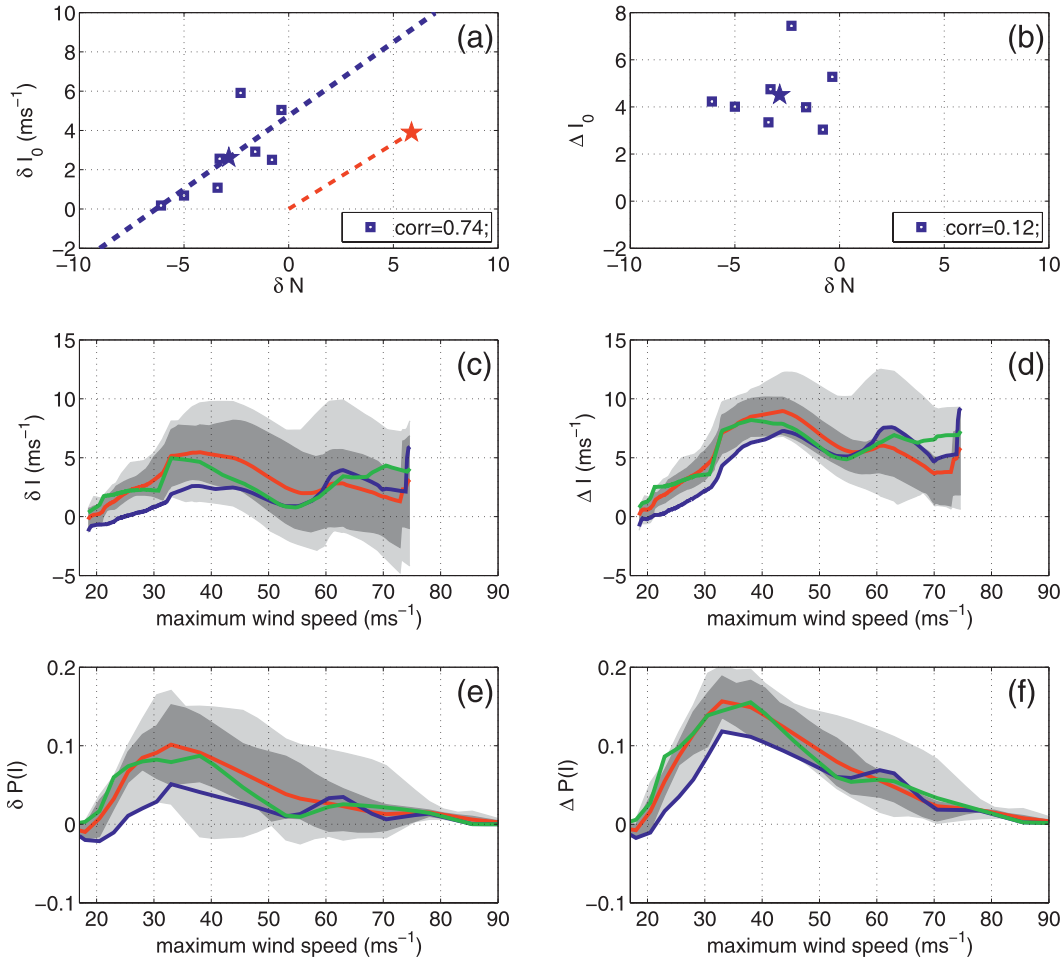

FIG. 7. (a) Scatterplot (squares) of the difference in TC mean intensity (I_0) and the difference in total number of TCs (N) for the North Atlantic between warmer and control experiments. The blue pentagram is the average across the eight experiments. The blue line indicates the linear regression. The red pentagram denotes the change of I_0 and N values between active and inactive years from 1981–2008 simulation. The red line connects the red pentagram to the origin, showing the slope. (b) As in (a), but δI_0 is replaced by ΔI_0. (c) Difference in intensity δI at given cumulative probabilities for the North Atlantic between warmer and control experiments. The abscissa is plotted in terms of intensity based on the control experiment $I(P)$. Red indicates the ensemble mean for the SST anomaly experiments; blue, the result for the 18-model ensemble mean SST anomaly experiment; and green, the result from the 2-K uniform warming experiments. Shaded areas show central 50% range and maximum and minimum changes. (d) As in (c), but for $\Delta I(I)$. (e) As in (c), but for the difference in cumulative probability $\delta P(I)$ for the North Atlantic between warmer and control experiments. (f) As in (e), but for $\Delta P(I)$.

climate, projections for changes in the frequency of intense hurricanes are difficult. Indeed, Figs. 6b and 6c suggest that the spread among models in the changes in frequency of intense storms comes from the spread in both δN and δP (note that the scale in the δP plot is half of that in other plots in this figure.) This spread in δP is solely due to the differences in the SST warming patterns that also result in the change in δN. This suggests an alternative way of decomposing the total response into pieces dominated by different processes.

Consider the mean over all TCs of the lifetime maximum intensity, I_0, a quantity already discussed above in the context of interannual variability. Figure 7a shows a scatterplot of δI_0 and δN for the North Atlantic obtained by differencing each global warming experiment with the control. There is an approximately linear relationship between the two with a correlation coefficient of about 0.7. The linear regression yields a slope of roughly 0.8 m s^{-1} per TC across the eight experiments. This slope is close to the value of $\delta I_0/\delta N$ estimated from the active and inactive years for the 1981–2008 simulation (red star in Fig. 7a). This common slope suggests some commonality to the dynamics controlling the covariation of intensity and total TC frequency. However, the average intensity across the

eight experiments increases despite a substantial reduction of the mean N in the warmer climate.

We are led to the picture that there are a set of environmental factors (e.g., SST indices) that control both total TC frequency and statistical measures of intensity in both the interannual and global warming simulations, but that there is another factor x that also comes into play when considering changes in intensity due to global warming, causing I_0 to increase. The change of I_0 with global warming may then be written as

$$\delta I_0 = \frac{\partial I_0}{\partial N}\delta N + \frac{\partial I_0}{\partial x}\delta x. \quad (3)$$

Figure 7b shows the same plot as in Fig. 7a but with δI_0 replaced by

$$\Delta I_0 \equiv \delta I_0 - \frac{\partial I_0}{\partial N}\delta N, \quad (4)$$

where the slope in the second term is taken from the active/inactive composite (the dotted red line in Fig. 7a). With roughly 50% of the variance removed, all the models now exhibit a rise of intensity with a mean increase of 4.5 m s^{-1} and a standard deviation of 1.4 m s^{-1} across the models. This represents the intensity change adjusted to remove the part of the response that is evidently associated with the change in N. One would expect this intrinsic intensity change to be roughly consistent with the estimate of intensity response in simulations of single storms in idealized environments perturbed to correspond to global warming conditions (Knutson and Tuleya 2004).

The intermodel spread in δI_0 is also correlated (coefficient = 0.78) with the model-simulated change in seasonal (August–October) mean Atlantic MDR maximum potential intensity (MPI), a proposed theoretical upper limit for TC intensity (Emanuel 1988). Seven out of the eight warming experiments produce an increase of MPI in the Atlantic MDR. (MPI is computed using the ratio of thermal to mechanical drag coefficient of 0.9 and assuming a reversible adiabatic ascent.) The linear regression of δMPI and δN yields somewhat a smaller slope (~0.6 m s^{-1} per TC) than $\delta I_0/\delta N$ (~0.8 m s^{-1} per TC). This is also true for the historical simulations for the active and inactive years. When the δN-dependent part of δMPI is removed (as in Fig. 7b), the mean change of MPI is roughly 2.5 m s^{-1}. This is substantially smaller than that of ΔI_0. While it is plausible that the seasonally and spatially averaged MPI may underestimate the actual MPI of individual TCs, this difference also suggests that besides the change of MPI itself—the change of the shape of the intensity probability distribution toward more intense storms—may play an important role in determining the simulated change of TC mean intensity in a warmer climate.

A more detailed view of the TC intensity response can be obtained from the intensity change $\delta I(I)$ as defined in section 2. Analogous to the procedure used above, we can define $\Delta I(I)$ by removing the δN-dependent part using the active/inactive composite for $\delta I(I)/\delta N$. Most of the models (Fig. 7c) produce positive $\delta I(I)$ for all values of I, but there is large spread across the models, ranging from small negative values to positive values as large as +10 m s^{-1}. Plotting $\Delta I(I)$ considerably reduces the model spread, with all models now displaying clear positive $\Delta I(I)$ for all values of intensity. For hurricanes up to category 3, the increase is roughly 5–10 m s^{-1} whereas for category 4 and higher intensities ($I \geq 60$ m s^{-1}), the model spread becomes somewhat larger, ranging from 3 to 12 m s^{-1}. The smaller change in I_0 is due to the large weight in the averaging given to the large number of weak storms for which the intensity changes are relatively small. [The sharp change in $\Delta I(I)$ at the hurricane threshold, 33 m s^{-1}, is due to our bilinear statistical adjustment of intensities, with the adjustment increasing rapidly in magnitude precisely at this value of I. Smoothing out this adjustment would result in a smoother plot of $\delta I(I)$.]

As a final perspective, we also plot $\delta P(I)$ and $\Delta P(I)$ in this figure. Once again, the spread in $\delta P(I)$ is reduced when computing $\Delta P(I)$, with the peak value of the latter close to 0.15 at the hurricane threshold. Since these are changes in cumulative probability, the derivative with respect to I corresponds to changes in the probability density function of storm intensity. The distribution of $\Delta P(I)$ indicates, after subtracting out the "N dependence," that there is a relatively continuous increase in the probability of storms at intensities above 40 m s^{-1} with a fairly sharp reduction in the probability of storms below hurricane intensity. For category 4 and stronger hurricanes ($I \geq 60$ m s^{-1}), ΔP is roughly 0.03–0.13, corresponding to a 20%–90% increase in probability, taking the statistical adjustment at face value. Some of this large spread is likely due to sampling noise resulting from the small number of storms involved in these estimates. The effects of changes in δN must be added to this intrinsic intensity increase associated with global warming.

4. Discussion

We have explored the implications of using a statistical adjustment to the intensity distribution of tropical cyclones generated by the HIRAM GCM running at approximately 50-km resolution. The hurricane frequencies in the North Atlantic simulated by this model, running over observed SSTs (HADISST1.1) from 1981

to 2008, appear to be realistic as judged by simulations of interannual variability and trend. But the model does not simulate any hurricanes with maximum wind speeds greater than 50 m s^{-1} (category 3 or higher) in the North Atlantic. Using a simple statistical adjustment that is constructed so as to generate a realistic distribution of intensities, the resulting model captures much of the intensity change between active and inactive years in the North Atlantic over this historical period.

The observations used to test the model are composites of active and inactive years. Three different composites are utilized: the 10 most active versus 10 least active years, El Niño versus La Niña years, and the first half of this period versus the second half. Even though they are not independent, it is reassuring that there is in each case a similar clear increase in measures of storm intensity in active years as compared to inactive years. A test of this kind is valuable, if not essential, to help judge the value of dynamical or statistical models for addressing the question of global warming and hurricane intensity. What the best metric is for this comparison of model and observations, given the noise in intensity distributions within individual years, remains an important open question in our view.

Global warming simulations are conducted by using SST anomalies generated for the A1B scenario over the twenty-first century by a suite of coupled models for which results are archived in the CMIP3 database. An ensemble SST anomaly is also considered, as well as an idealized uniform increase of 2 K. The CO_2 concentration is also doubled in each of these simulations.

As discussed with a smaller ensemble of simulations in ZHLV, on average there is a tendency for a reduction in the number of tropical cyclones in the North Atlantic, but there is a large spread when using the different SST anomalies from different coupled models. This spread is due to differences in spatial structure of the SST anomaly; specifically, SSTs from models that warm the tropical Atlantic less than the tropics as a whole generate larger decreases in TC frequency than the average model, whereas SSTs from models that warm the Atlantic more than the tropical mean generate a smaller decrease, or even an increase. After statistical adjustment, we find an analogous result for various measures of mean intensity: various statistical measures of the change in intensity are positively correlated across the ensemble with the change in frequency, just as for interannual variability in the historical period that we use for model testing.

Based on this result, we examine the consequences of adjusting these intensity changes for the differing changes in frequency, assuming that the mechanisms causing the frequency/intensity correlation in observations (and in our historical simulations) are active in the global warming simulations as well. This manipulation leaves behind residual intensity changes that occur in the global warming simulations that have no direct analog in interannual variability. This residual, or intrinsic, intensity response has reduced intermodel spread. We find roughly a 5–10 m s^{-1} increase in intensity for storms in the 30–60 m s^{-1} range in these twenty-first-century projections. This intrinsic intensity shift, which is not overly sensitive to the spatial structure of the warming, is then modified by the part of the response that is correlated with the frequency change, which is strongly dependent on spatial structure.

A plausible explanation for this behavior would be that some factor closely related to potential intensity (Emanuel 1988) is influencing the global warming simulations in a manner that does not come into play in the active/inactive composites (or our simulations of those observations). One can similarly decompose the potential intensity into two components, one that is related to frequency and one that is not. When the frequency-dependent component of PI response is removed, we find an intrinsic increase of PI about 2–3 m s^{-1}, which is close to the PI change generated from the uniform 2-K warming experiment. However, this change in seasonally and spatially averaged potential intensity alone appears not to be sufficient to explain the simulated mean intensity change (4–5 m s^{-1}), indicating that changes in the shape of the intensity distribution may also be important. One might also consider the possibility of direct effects of CO_2 as relevant in this regard, but initial analysis of experiments with CO_2 increase and unmodified SSTs (to be reported elsewhere) suggests that this direct effect is a weakening, rather than a strengthening, of intensities in this model.

All of our conclusions concerning intensity rely on the statistical intensity adjustment that we utilize. Dynamical approaches are ultimately more convincing. Using a dynamical downscaling approach, Bender et al. (2010), for example, suggests a transition point from a reduction to an increase at intensity between categories 3 and 4 (50–60 m s^{-1}) using an ensemble mean SST projection. However, our results from the ensemble mean SST indicates a transition point at a higher intensity (60–70 m s^{-1}). Moreover, only half of the projected SSTs produce an overall increase of frequency at intensity around 60 m s^{-1} in our simulation. We emphasize that because of the potential competing effects between changes in total storm frequency (δN) and changes in intensity probability distribution (δP) in a greenhouse gas–warmed climate, model errors in predicting both δN and δP will impact the overall projection of intense hurricanes. For example, the modeling framework of Bender et al. (2010) appears to underestimate the intensity difference between active and inactive decades in its historical simulations. Since it

also projects a reduction of TCs on average, the suggestion would be that it might be overestimating the intensity increase with global warming by underestimating the reduction in intensity associated with the reduction in frequency. Further comparisons with this and other modeling approaches [such as that of Emanuel et al. (2008)] should be informative, especially if the results from these different methodologies for active/inactive composites from the historical record are documented.

Acknowledgments. This research used resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the Department of Energy under Contract DE-AC05-00OR22725. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment Program. The authors thank G. Vecchi, S. Garner, and T. Knutson for valuable discussions and S.-J. Lin for his leading role in developing this model. Ming Zhao was supported in part by NSF Grant ATM-0612551 and in part under Award NA17RJ2612 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The findings are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration or the U.S. Department of Commerce.

APPENDIX

List of Active/Inactive and La Niña/El Niño Years

Ten most active years: 2005, 1995, 2008, 2003, 2004, 2001, 2000, 2007, 1998, 1996

Ten least active years: 1983, 1982, 1986, 1987, 1992, 1994, 1991, 1993, 1997, 2006

La Niña years: 1988, 1998, 1999, 2000, 2001, 2007

El Niño years: 1982, 1987, 1991, 1992, 1994, 1997, 2002

REFERENCES

Bender, M. A., T. R. Knutson, R. E. Tuleya, J. J. Sirutis, G. A. Vecchi, S. T. Garner, and I. M. Held, 2010: Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. *Science,* **327,** 454–458.

Bengtsson, L., K. I. Hodges, and M. Esch, 2007a: Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analyses. *Tellus,* **59A,** 396–416.

——, ——, ——, N. Keenlyside, L. Kornblueh, J.-J. Luo, and T. Yamagata, 2007b: How may tropical cyclones change in a warmer climate? *Tellus,* **59A,** 539–561.

Camargo, S., and A. Sobel, 2005: Western North Pacific tropical cyclones intensity and ENSO. *J. Climate,* **18,** 2996–3006.

Emanuel, K. A., 1988: The maximum intensity of hurricanes. *J. Atmos. Sci.,* **45,** 1143–1155.

——, 2000: A statistical analysis of tropical cyclone intensity. *Mon. Wea. Rev.,* **128,** 1139–1152.

——, R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.,* **89,** 347–367.

Gualdi, S., E. Scoccimarro, and A. Navarra, 2008: Changes in tropical cyclone activity due to global warming: Results from a high-resolution coupled general circulation model. *J. Climate,* **21,** 5204–5228.

Knutson, T. R., and R. E. Tuleya, 2004: Impact of CO_2-induced warming on simulated hurricane intensity and precipitation: Sensitivity to the choice of climate model and convective parameterization. *J. Climate,* **17,** 3477–3495.

——, and Coauthors, 2010: Tropical cyclones and climate change. *Nat. Geosci.,* **3,** 157–163, doi:10.1038/ngeo779.

Kruk, M. C., K. R. Knapp, D. H. Levinson, and J. Kossin, 2010: A technique for combining global tropical cyclone best track data. *J. Atmos. Oceanic Technol.,* **27,** 680–692.

LaRow, T., Y.-K. Lim, D. Shin, E. Chassignet, and S. Cocke, 2008: Atlantic basin seasonal hurricane simulations. *J. Climate,* **21,** 3191–3206.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007: The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.,* **88,** 1383–1394.

Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006: Tropical cyclone climatology in a global-warming climate as simulated in a 20 km mesh global atmospheric model: Frequency and wind intensity analyses. *J. Meteor. Soc. Japan,* **84,** 259–276.

Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *J. Comput. Phys.,* **227,** 55–78.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate,* **15,** 1609–1625.

Swanson, K. L., 2007: Impact of scaling behavior on tropical cyclone intensities. *Geophys. Res. Lett.,* **34,** L18815, doi:10.1029/2007GL030851.

——, 2008: Nonlocality of Atlantic tropical cyclone intensities. *Geochem. Geophys. Geosyst.,* **9,** Q04V01, doi:10.1029/2007GC001844.

Vecchi, G. A., K. L. Swanson, and B. J. Soden, 2008: Whither hurricane activity? *Nature,* **322,** 687–689.

Vitart, F., 2006: Seasonal forecasting of tropical storm frequency using a multi-model ensemble. *Quart. J. Roy. Meteor. Soc.,* **132,** 647–666.

Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate,* **22,** 6653–6678.

——, ——, and G. A. Vecchi, 2010: Retrospective forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Mon. Wea. Rev.,* **138,** 3858–3868.